

# Improved Recovery Bounds of Orthogonal Matching Pursuit using Restricted Isometry Property

Jian Wang and Byonghyo Shim

Information System Laboratory

School of Information and Communication, Korea University

Anam-dong, Seongbuk-gu, Seoul, Korea, 136-713

Email: {jwang,bshim}@isl.korea.ac.kr

**Abstract**—Orthogonal matching pursuit (OMP) is a greedy algorithm widely used for the recovery of sparse signals from compressed measurements. In this paper, we consider the perfect recovery of  $K$ -sparse signals using the OMP algorithm for two distinct scenarios, viz., when it performs  $K$  iterations and more than  $K$  iterations. In the first part of this paper, we show that if the sensing matrix satisfies the restricted isometry property (RIP) with  $\sqrt{K}\theta_{K,1} + \delta_K < 1$  then the OMP algorithm can perfectly recover any  $K$ -sparse signal in  $K$  iterations. Our result bridges the mutual incoherence property (MIP) and RIP based recovery bounds and also embraces many of recently proposed recovery bounds of the OMP. In the second part of this paper, we show that if the sensing matrix satisfies the RIP with  $\delta_{[8.93K]} < 0.03248$ , then the OMP perfectly recovers any  $K$ -sparse signal within  $6K$  iterations.

**Index Terms**—Sparse recovery, orthogonal matching pursuit (OMP), restricted isometric property (RIP), mutual incoherence property (MIP), compressive sensing (CS)

## I. INTRODUCTION

Recently there has been a growing interest in recovering sparse signals from compressed measurements [1]–[11]. The main goal of this task is to accurately estimate a high dimensional  $K$ -sparse vector  $\mathbf{x} \in \mathcal{R}^n$  ( $\|\mathbf{x}\|_0 \leq K$ ) from a small number of linear measurements  $\mathbf{y} \in \mathcal{R}^m$  ( $m \ll n$ ). The relationship between the signal vector and measurements is

$$\mathbf{y} = \Phi \mathbf{x} \quad (1)$$

where  $\Phi \in \mathcal{R}^{m \times n}$  is often called the sensing matrix. Since this system is underdetermined, infinitely many solutions exist and hence one cannot accurately recover the original signal vector  $\mathbf{x}$ . However, due to the prior information of signal sparsity, the signal vector  $\mathbf{x}$  can be accurately reconstructed via properly designed recovery algorithms. Among many recovery algorithms in the literature, orthogonal matching pursuit (OMP) algorithm has generated a great deal of interest in recent years for its competitive performance as well as practical benefits such as implementation simplicity and low computational complexity [5], [10], [12].

The OMP algorithm iteratively estimates the sparse signal  $\mathbf{x}$  and its support (i.e., index set of nonzero elements). Suppose the  $K$ -sparse vector  $\mathbf{x}$  is supported on  $T$  and let  $T^k$ ,  $\mathbf{x}^k$  and  $\mathbf{r}^k$  be the estimated support, the estimated sparse signal, and

This work was supported by the KCC (Korea Communications Commission), Korea, under the R&D program supervised by the KCA (Korea Communication Agency) (KCA-12-911-01-110) and the NRF grant funded by the Korea government (MEST) (No. 2012R1A2A2A01047510).

the residual in the  $k$ -th iteration, respectively. Then the OMP repeats the following operations until the residual equals zero or the iteration number reaches a predefined maximum value:

- Among correlations between  $\phi_i$  ( $i$ -th column of  $\Phi$ ) and the residual  $\mathbf{r}^{k-1}$  generated in the  $(k-1)$ -th iteration, find the largest element in magnitude and the corresponding index  $t^k$ . That is,

$$t^k = \arg \max_{i \in \{1, \dots, n\} \setminus T^{k-1}} |\langle \mathbf{r}^{k-1}, \phi_i \rangle|. \quad (2)$$

- Once  $t^k$  is identified, add this index into the estimated support set

$$T^k = T^{k-1} \cup \{t^k\}. \quad (3)$$

- Solve the least squares (LS) problem

$$\mathbf{x}^k = \arg \min_{\text{supp}(\mathbf{v})=T^k} \|\mathbf{y} - \Phi \mathbf{v}\|_2. \quad (4)$$

- Update the residual of the  $k$ -th iteration as

$$\mathbf{r}^k = \mathbf{y} - \Phi \mathbf{x}^k. \quad (5)$$

Readers are referred to [12] for more details.

The OMP algorithm has long been considered as a heuristic algorithm hard to be analyzed. Recently, however, many efforts have been made to discover conditions of the OMP ensuring the perfect recovery of sparse signals. In the analysis, a condition so called the mutual incoherence property (MIP) has been employed [1]. The mutual coherence parameter  $\mu(\Phi)$  of the sensing matrix  $\Phi$  is defined as

$$\mu(\Phi) = \max_{i \neq j} |\langle \phi_i, \phi_j \rangle|$$

where  $\phi_i$  and  $\phi_j$  are two distinct column vectors of  $\Phi$ . Tropp showed that if  $\Phi$  has unit  $\ell_2$ -norm columns and satisfies  $\mu(\Phi) < \frac{1}{2K-1}$ , then the OMP algorithm can reconstruct  $K$ -sparse signals accurately in  $K$  iterations [12]. An alternate and recently popular framework is the restricted isometric property (RIP) [2]. A sensing matrix  $\Phi$  is said to satisfy the RIP of order  $K$  if there exists a constant  $\delta(\Phi)$  such that

$$(1 - \delta(\Phi)) \|\mathbf{x}\|_2^2 \leq \|\Phi \mathbf{x}\|_2^2 \leq (1 + \delta(\Phi)) \|\mathbf{x}\|_2^2 \quad (6)$$

for any  $K$ -sparse vector  $\mathbf{x}$ . In particular, the minimum of all constants  $\delta(\Phi)$  satisfying (6) is called the isometric constant  $\delta_K(\Phi)$ . Also, as a related quantity, the  $(K, K')$ -restricted orthogonality constant  $\theta_{K, K'}(\Phi)$  is defined as the minimum value satisfying

$$|\langle \Phi \mathbf{x}, \Phi \mathbf{x}' \rangle| \leq \theta_{K, K'}(\Phi) \|\mathbf{x}\|_2 \|\mathbf{x}'\|_2 \quad (7)$$

for all  $K$  and  $K'$ -sparse vectors  $\mathbf{x}$  and  $\mathbf{x}'$  having disjoint supports and  $K + K' \leq n$ . In the sequel, we use  $\mu$ ,  $\delta_K$ , and  $\theta_{K,K'}$  instead of  $\mu(\Phi)$ ,  $\delta_K(\Phi)$ , and  $\theta_{K,K'}(\Phi)$  for notational simplicity. Davenport and Wakin showed that, under  $\delta_{K+1} < \frac{1}{3\sqrt{K}}$ , the OMP algorithm guarantees the perfect reconstruction of any  $K$ -sparse signal in  $K$  iterations [10]. Liu and Temlyakov improved the condition to  $\delta_{K+1} < \frac{1}{\sqrt{2K}}$  [13], and Wang and Shim [14] and Mo and Shen [15] further improved the condition to  $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$ . They showed the optimality of the condition by providing an example that there exist  $K$ -sparse signals which cannot be recovered under  $\delta_{K+1} = \frac{1}{\sqrt{K}}$ .

While many studies on the OMP algorithm has focused on the case where the iteration number is limited to  $K$ , there has been some works investigating the behavior of the OMP algorithm when it performs more than  $K$  iterations [16]–[18] or when it chooses more than one indices per iteration [19]. Both in theoretical performance guarantees and empirical simulations, these approaches provide better results and also offers new insights into the algorithm. Livshitz showed that  $\delta_{\alpha K^{1.2}} = \beta K^{-0.2}$  guarantees the exact reconstruction of  $K$ -sparse signals within  $\lfloor \alpha K^{1.2} \rfloor$  iterations [16]. The main benefit of this result is that the measurement size ensuring the perfect recovery is  $m = O(K^{1.6} \log n)$ , which is clearly less than  $m = O(K^2 \log n)$  for the OMP with  $K$  iterations [10], [13], [14]. Unfortunately, since  $\alpha \sim 10^5$  and  $\beta \sim 10^{-6}$ , it is not easy to enjoy the benefit in practice. Recently, it has been shown by Zhang that  $\delta_{31K} < \frac{1}{3}$  guarantees the perfect recovery of  $K$ -sparse signals within  $30K$  iterations [17] and this result has been improved by Foucart to  $\delta_{22K} < \frac{1}{6}$  with  $12K$  maximal iterations [18]. The moral of the story in these works is that the OMP can actually recover  $K$ -sparse signals in  $cK$  ( $c > 1$ ) iterations when  $\delta_{O(K)}$  is upper bounded by a constant, which in turn implies that the required measurement size is  $m = O(K \log n)$ .

In this paper, we study the perfect recovery condition of the OMP algorithm for two distinct scenarios: when the OMP performs 1) strictly  $K$  iterations and 2) more than  $K$  iterations. We henceforth refer to these two schemes as  $\text{OMP}_K$  and  $\text{OMP}_{cK}$ , respectively. The main contributions of this paper are twofold:

- In the first part of this paper, we provide the RIP condition for the  $\text{OMP}_K$ , expressed in terms of  $\delta_K$  and  $\theta_{K,K'}$ . We show that the perfect recovery of any  $K$ -sparse signal can be ensured if the sensing matrix  $\Phi$  satisfies  $\sqrt{K}\theta_{K,1} + \delta_K < 1$ . We also show that this condition is optimal in the sense that the perfect recovery is violated even with the slight relaxation of the proposed condition. Interestingly, the proposed result embraces many of recently proposed bounds of the  $\text{OMP}_K$ . For example:
  - By combining the well known “ $\delta_K - \mu$ ” inequality  $\delta_K \leq (K-1)\mu$  [6], [12], [20] and the proposed condition, we obtain the MIP based recovery bound  $\mu < \frac{1}{2K-1}$  [12] (**Theorem 2.6**).
  - Using the proposed condition together with the simple relationship between  $\theta_{K,K'}$  and  $\delta_K$  ( $\theta_{K,K'} \leq \delta_{K+K'}$  [2]), we obtain the RIP based recovery bound

$$\delta_{K+1} < \frac{1}{\sqrt{K+1}} \quad [14] \quad (\textbf{Theorem 2.7}).$$

- When combined with the square root lifting inequality  $\theta_{\omega K, K'} \leq \sqrt{\omega}\theta_{K, K'}$  [9], we achieve the RIP based recovery bound  $\delta_K < \frac{\sqrt{K-1}}{\sqrt{K-1}+K}$  [21] (**Theorem 2.8**).

- As mentioned in [17], [18], the  $\text{OMP}_{cK}$  has less stringent and hence better recovery condition compared to the condition of  $\text{OMP}_K$ . In the second part of this paper, we establish a new recovery condition for the  $\text{OMP}_{cK}$ . To be specific, we show that if the sensing matrix  $\Phi$  satisfies the RIP with  $\delta_{[8.93K]} < 0.03248$ , then it is possible to recover any  $K$ -sparse signal within  $6K$  iterations using the  $\text{OMP}_{cK}$  algorithm.

The rest of this paper is organized as follows: In Section II, we present an optimal RIP condition for the  $\text{OMP}_K$  to achieve the perfect recovery of  $K$ -sparse signals. In Section III, we provide the perfect recovery condition for the  $\text{OMP}_{cK}$ . We conclude our paper in Section IV.

We summarize notations that will be used in this paper.  $T = \text{supp}(\mathbf{x}) = \{i | x_i \neq 0\}$  is the set of non-zero positions in  $\mathbf{x}$ .  $\Omega = \{1, 2, \dots, n\}$ . For  $S \subseteq \Omega$ ,  $|S|$  is the cardinality of  $S$ .  $T \setminus S$  is the set of all elements contained in  $T$  but not in  $S$ .  $\Phi_S \in \mathcal{R}^{m \times |S|}$  is a submatrix of  $\Phi$  that only contains columns indexed by  $S$ .  $\Phi'_S$  is the transpose of the matrix  $\Phi_S$ .  $\mathbf{x}_S \in \mathcal{R}^{|S|}$  is an vector which equals  $\mathbf{x}$  for elements indexed by  $S$ . If  $\Phi_S$  is full column rank, then  $\Phi_S^\dagger = (\Phi'_S \Phi_S)^{-1} \Phi'_S$  is the pseudoinverse of  $\Phi_S$ .  $\text{span}(\Phi_S)$  is the span of columns in  $\Phi_S$ .  $\mathbf{P}_S = \Phi_S \Phi_S^\dagger$  is the projection onto  $\text{span}(\Phi_S)$ .  $\mathbf{P}_S^\perp = \mathbf{I} - \mathbf{P}_S$  is the projection onto the orthogonal complement of  $\text{span}(\Phi_S)$ .

## II. PERFECT RECONSTRUCTION OF $K$ -SPARSE SIGNALS VIA $\text{OMP}_K$

### A. Sufficient Condition

The main result for the  $\text{OMP}_K$  is given in the following theorem.

*Theorem 2.1:* Let  $\mathbf{x} \in \mathcal{R}^n$  be any  $K$ -sparse vector with support  $T$  and let  $\Phi \in \mathcal{R}^{m \times n}$  be the sensing matrix satisfying

$$\sqrt{K}\theta_{K,1} + \delta_K < 1, \quad (8)$$

then the  $\text{OMP}_K$  perfectly recovers  $\mathbf{x}$  from  $\mathbf{y} = \Phi\mathbf{x}$ ,

Before we proceed, we provide lemmas useful in our analysis.

*Lemma 2.2 (Monotonicity property [2]):* If a sensing matrix  $\Phi$  satisfies the RIP of both orders  $K$  and  $K'$ , then  $\delta_K \leq \delta_{K'}$  for any  $K \leq K'$ .

*Lemma 2.3 (A direct consequence of RIP [22]):* Let  $I \subset \Omega$  and  $\Phi_I$  be the restriction of the columns of  $\Phi$  indexed by  $I$ . If  $\delta_{|I|} < 1$ , then for any vector  $\mathbf{x} \in \mathcal{R}^{|I|}$ ,

$$(1 - \delta_{|I|})\|\mathbf{x}\|_2 \leq \|\Phi'_I \Phi_I \mathbf{x}\|_2 \leq (1 + \delta_{|I|})\|\mathbf{x}\|_2.$$

*Lemma 2.4 (Square root lifting inequality [9]):* For any  $\omega \geq 1$  and positive integers  $K, K'$  such that  $\omega K'$  is an integer,  $\theta_{\omega K, K'} \leq \sqrt{\omega}\theta_{K, K'}$ .

*Lemma 2.5:* Let  $I_1, I_2 \in \Omega$  be two disjoint sets ( $I_1 \cap I_2 = \emptyset$ ). Then for any vector  $\mathbf{x}$  supported on  $I_2$ ,

$$\|\Phi'_{I_1} \Phi_{I_2} \mathbf{x}_{I_2}\|_2 \leq \theta_{|I_1|, |I_2|} \|\mathbf{x}\|_2. \quad (9)$$

*Proof:* See Appendix A.  $\blacksquare$

### Proof of Theorem 2.1

*Proof:* We will prove the theorem using mathematical induction. In the first iteration ( $k = 1$ ) of the  $\text{OMP}_K$ , an index  $t^1$  of the column maximally correlated with the measurement  $\mathbf{y}$  is chosen. That is,  $t^1 = \arg \max_i |\langle \phi_i, \mathbf{y} \rangle|$ . Then, we have

$$|\langle \phi_{t^1}, \mathbf{y} \rangle| = \|\Phi' \mathbf{y}\|_\infty \quad (10)$$

$$\geq \|\Phi'_T \mathbf{y}\|_\infty \quad (11)$$

$$\geq \frac{1}{\sqrt{K}} \|\Phi'_T \mathbf{y}\|_2, \quad (12)$$

where (12) is due to  $|T| = K$ . Further, since  $\mathbf{y} = \Phi_T \mathbf{x}_T$ , we have

$$\begin{aligned} |\langle \phi_{t^1}, \mathbf{y} \rangle| &\geq \frac{1}{\sqrt{K}} \|\Phi'_T \Phi_T \mathbf{x}_T\|_2 \\ &\geq \frac{1}{\sqrt{K}} (1 - \delta_K) \|\mathbf{x}_T\|_2, \end{aligned} \quad (13)$$

where (13) follows from Lemma 2.3.

Now, suppose that  $t^1$  does not belong to the support of  $\mathbf{x}$ . Then  $t^1 \cap T = \emptyset$  and hence from Lemma 2.5, we have

$$\begin{aligned} |\langle \phi_{t^1}, \mathbf{y} \rangle| &= \|\phi'_{t^1} \Phi_T \mathbf{x}_T\|_2 \\ &\leq \theta_{1,K} \|\mathbf{x}_T\|_2. \end{aligned} \quad (14)$$

This case, however, will never be true if

$$\frac{1}{\sqrt{K}} (1 - \delta_K) \|\mathbf{x}_T\|_2 > \theta_{1,K} \|\mathbf{x}_T\|_2.$$

Equivalently,

$$\sqrt{K} \theta_{1,K} + \delta_K < 1.$$

In summary, if  $\sqrt{K} \theta_{1,K} + \delta_K < 1$ , then  $t^1 \in T$  for the first iteration of the  $\text{OMP}_K$ .

Now we consider the general (non-initial) iteration of the  $\text{OMP}_K$ . In this step, we show that under the hypothesis that the former  $k$  iterations are successful, the  $(k+1)$ -th iteration should also be successful. To be specific, we show that if  $T^k = \{t^1, t^2, \dots, t^k\} \subset T$ , then  $t^{k+1}$  is in  $T$  but not in  $T^k$  ( $t^{k+1} \in T \setminus T^k$ ).

First, recall that the residual of the  $k$ -th iteration is given by

$$\mathbf{r}^k = \mathbf{y} - \Phi \mathbf{x}^k. \quad (15)$$

Since  $\text{supp}(\mathbf{x}^k) = T^k \subset T$  by the hypothesis and  $\text{supp}(\mathbf{x}) = T$ , one can easily deduce that  $\mathbf{r}^k$  is a linear combination of the  $|T| (= K)$  columns of  $\Phi_T$ . Accordingly, there exists a  $K$ -sparse signal  $\Delta = \mathbf{x} - \mathbf{x}^k$  supported on  $T$  such that

$$\mathbf{r}^k = \Phi_T \Delta_T = \Phi \Delta. \quad (16)$$

Clearly,  $\mathbf{r}^k$  is the measurements of the  $K$ -sparse signal  $\Delta$  using the sensing matrix  $\Phi$ . Since  $\mathbf{r}^k$  preserves the sparsity level of the original signal  $\mathbf{x}$ , if  $\sqrt{K} \theta_{1,K} + \delta_K < 1$ , the index  $t^{k+1}$  chosen in the  $(k+1)$ -th iteration belongs to the support of  $\Delta$  (i.e.,  $t^{k+1} \in T$ ). Recalling that the already selected index will not be re-selected (see Eq. (2)),  $t^{k+1} \in T \setminus T^k$ . This concludes the proof.  $\blacksquare$

### B. Optimality

Theorem 2.1 characterizes the “success” of the  $\text{OMP}_K$  using the isometric constant  $\delta_K$  and the restricted orthogonality constant  $\theta_{K,K'}$ . These two quantities are related respectively to the mutual coherence parameter  $\mu$  by [6], [12], [20]

$$\delta_K \leq (K-1) \mu \quad (17)$$

and [23]

$$\theta_{K,K'} \leq \sqrt{KK'} \mu. \quad (18)$$

By a simple connection between these inequalities and Theorem 2.1, one can obtain the Tropp’s MIP condition [12].

*Theorem 2.6 (Alternative proof of  $\mu < \frac{1}{2K-1}$  [12]):* Let  $\Phi$  be the sensing matrix with unit  $\ell_2$ -norm columns. Then under  $\mu < \frac{1}{2K-1}$ , the  $\text{OMP}_K$  perfectly recovers any  $K$ -sparse signal  $\mathbf{x}$  from the measurements  $\mathbf{y} = \Phi \mathbf{x}$ .

*Proof:* Plugging  $\delta_K \leq (K-1)\mu$  and  $\theta_{K,1} \leq \sqrt{K}\mu$  into  $\sqrt{K}\theta_{K,1} + \delta_K < 1$ , we have

$$K\mu + (K-1)\mu < 1, \quad (19)$$

or

$$\mu < \frac{1}{2K-1}. \quad (20)$$

The optimality of the proposed recovery condition in (8) is justified by showing that the perfect recovery is violated even with the slight relaxation of (8). To be specific, due to the use of inequalities (17) and (18) in the proof of Theorem 2.6, the proposed condition is essentially guaranteed by (20) and hence better than (less stringent) or at least equivalent to (20).<sup>1</sup> Indeed, considering the optimality of condition (20),<sup>2</sup> we claim that the proposed condition is optimal.

### C. Relations to the previous RIP bounds

The proposed condition consists of two quantities: the isometric constant  $\delta_K$  and the restricted orthogonality constant  $\theta_{K,K'}$ . In order to obtain the condition expressed in terms of the isometric constant only, we use [2]

$$\theta_{K,K'} \leq \delta_{K+K'}. \quad (21)$$

Combining (21) and (8), we obtain the RIP based recovery condition [12].

*Theorem 2.7 (Alternative proof of  $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$  [14]):* The  $\text{OMP}_K$  recovers any  $K$ -sparse signal  $\mathbf{x}$  from the measurement  $\Phi \mathbf{x}$  if the sensing matrix  $\Phi$  satisfies

$$\delta_{K+1} < \frac{1}{\sqrt{K+1}}. \quad (22)$$

*Proof:* Applying  $K' = 1$  in (21), we have

$$\theta_{K,1} \leq \delta_{K+1}. \quad (23)$$

<sup>1</sup>When the bound becomes less stringent, the set of feasible sensing matrices  $\Phi$  for which the perfect recovery is possible gets larger.

<sup>2</sup>The optimality of condition (20) has been proven in [23, Theorem 3.1] by showing that the  $\text{OMP}_K$  can fail to recover  $K$ -sparse signals under  $\mu = \frac{1}{2K-1}$ .

TABLE I  
RECOVERY BOUNDS OF THE  $\text{OMP}_K$  ALGORITHM

Used property	Recovery bound	Remarks
MIP	$\mu < \frac{1}{2K-1}$ [12]	optimal
RIP	$\sqrt{K}\theta_{K,1} + \delta_K < 1$	optimal
	$\delta_K < \frac{\sqrt{K-1}}{\sqrt{K-1}+K}$ [21]	near optimal
	$\delta_{K+1} < \frac{1}{\sqrt{K+1}}$ [14]	near optimal

Using the monotonicity property of the isometry constant (Lemma 2.2), we have

$$\delta_K \leq \delta_{K+1}. \quad (24)$$

The theorem is established by plugging (23) and (24) into (8).  $\blacksquare$

Note that the recovery condition in Theorem 2.7 can be obtained by imposing two relaxations in (23) and (24) on Theorem 2.1. For sure, when a different relaxation is applied, one can obtain the different recovery condition. Indeed, with the help of the square root lifting inequality  $\theta_{\omega K, K'} \leq \sqrt{\omega}\theta_{K, K'}$  [9], one can obtain the RIP condition expressed in terms of  $\delta_K$ , which is presumably smallest order of the isometry constant and hence most natural and easy to interpret.

*Theorem 2.8 (Alternative proof of  $\delta_K < \frac{\sqrt{K-1}}{\sqrt{K-1}+K}$  [21]):* The  $\text{OMP}_K$  recovers any  $K$ -sparse ( $K > 1$ ) signal  $\mathbf{x}$  from the measurements  $\mathbf{y} = \Phi\mathbf{x}$  if the sensing matrix  $\Phi$  satisfies

$$\delta_K < \frac{\sqrt{K-1}}{\sqrt{K-1}+K}. \quad (25)$$

*Proof:* Let  $\omega = \frac{K}{K-1}$ , then clearly  $\omega \geq 1$ . Using the square root lifting inequality in Lemma 2.4, we have

$$\theta_{K,1} = \theta_{\omega(K-1),1} \quad (26)$$

$$\leq \sqrt{\omega}\theta_{K-1,1} \quad (27)$$

$$\leq \sqrt{\omega}\delta_K \quad (28)$$

$$= \sqrt{\frac{K}{K-1}}\delta_K, \quad (29)$$

where (28) is because  $\theta_{K,K'} \leq \delta_{K+K'}$ . Using (8) and (29), one can easily show that (25) is the sufficient condition for the perfect recovery of the  $K$ -sparse signal  $\mathbf{x}$ .  $\blacksquare$

In Table I, we summarize the recovery bounds of the  $\text{OMP}_K$  algorithm. Note that among four bounds in the table, the proposed bound is the weakest.

### III. PERFECT RECONSTRUCTION OF $K$ -SPARSE SIGNALS VIA $\text{OMP}_{cK}$

In this section, we analyze the RIP based condition under which  $\text{OMP}_{cK}$  can perfectly recover any  $K$ -sparse signal. The  $\text{OMP}_{cK}$  can recover  $K$ -sparse signals accurately when all indices in the support are chosen within  $cK$  iterations. Therefore, when the iteration loop of the  $\text{OMP}_{cK}$  is finished, the finally selected index set  $T^f$  (where  $f$  denotes the final iteration number) may contain indices not in  $T$ . Even in this

situation, the final result is unaffected and the original signal  $\mathbf{x}$  is perfectly recovered as long as  $T \subseteq T^f$  ( $f \leq m$ ) because<sup>3</sup>

$$\mathbf{x}^f = \arg \min_{\text{supp}(\mathbf{x})=T^f} \|\mathbf{y} - \Phi\mathbf{x}\|_2 \quad (30)$$

and

$$\begin{aligned} (\mathbf{x}^f)_{T^f} &= \Phi_{T^f}^\dagger \mathbf{y} \\ &= \Phi_{T^f}^\dagger \Phi_{T^f} \mathbf{x}_{T^f} \\ &= \Phi_{T^f}^\dagger \Phi_{T^f} \mathbf{x}_{T^f} - \Phi_{T^f}^\dagger \Phi_{T^f \setminus T} \mathbf{x}_{T^f \setminus T} \\ &= \mathbf{x}_{T^f}, \end{aligned} \quad (31)$$

where (31) follows from the fact that  $\mathbf{x}_{T^f \setminus T} = \mathbf{0}$ .

As mentioned, it has been shown in [17], [18] that  $T \subseteq T^{30K}$  and  $T \subseteq T^{12K}$  under the condition  $\delta_{31K} < \frac{1}{3}$  and  $\delta_{22K} < \frac{1}{6}$ , respectively.<sup>4</sup> Unfortunately, these conditions impose a bit too many iterations and also cause a limitation on the sparsity level  $K$ . For example,  $12K$  iterations require the sensing matrix to be at least  $12K$  rows ( $m \geq 12K$ ), which in turn requires  $K \leq m/12$ . In short, our main result is that  $T \subseteq T^{6K}$  holds true under  $\delta_{[8.93K]} < 0.03248$ . Roughly speaking, our result indicates that the  $\text{OMP}_{cK}$  can perform accurate reconstruction of  $K$ -sparse signals with only half number of iterations of what has been proved thus far. Our main result is formally described in the following theorem.

*Theorem 3.1:* Let  $\mathbf{x} \in \mathcal{R}^n$  be any  $K$ -sparse signal supported on  $T$  and  $\Phi \in \mathcal{R}^{m \times n}$  be the sensing matrix with unit  $\ell_2$ -norm columns. Further, let  $\Gamma^k = T \setminus T^k$  be the set of remaining support indices after  $k$  iterations of the  $\text{OMP}_{cK}$ . If  $\delta_{[8.93K]} < 0.03248$ , then  $T \subseteq T^{k+6|\Gamma^k|}$ .

The key point of the theorem is that after  $k$  iterations of the  $\text{OMP}_{cK}$ , the remaining support indices in  $\Gamma^k$  will be selected within  $6|\Gamma^k|$  additional iterations as long as  $\delta_{[8.93K]} < 0.03248$ .

*Corollary 3.2:* Let  $\Phi \in \mathcal{R}^{m \times n}$  be the sensing matrix having unit  $\ell_2$ -norm columns and let  $\mathbf{x} \in \mathcal{R}^n$  be any  $K$ -sparse vector with support  $T$ . Then, under  $\delta_{[8.93K]} < 0.03248$ , the  $\text{OMP}_{cK}$  perfectly recovers  $\mathbf{x}$  from  $\mathbf{y} = \Phi\mathbf{x}$  within  $6K$  iterations.

*Proof:* Applying  $k = 0$  in Theorem 3.1,  $\Gamma^k = T \setminus T^0 = T$  and hence  $T^{k+6|\Gamma^k|} = T^{6|T|} = T^{6K} \supseteq T$ .  $\blacksquare$

### A. Preliminaries

Before presenting the proof of Theorem 3.1, we provide some useful definitions, lemmas, and propositions. Without loss of generality, we assume that  $\Gamma^k = \{1, 2, \dots, |\Gamma^k|\}$ . Then it is clear that  $1 \leq |\Gamma^k| \leq K$ . For example, if  $k = 0$ , then  $T^k = \emptyset$  and  $|\Gamma^k| = |T| = K$ . Whereas if  $T^k \supseteq T$ , then  $\Gamma^k = \emptyset$  and  $|\Gamma^k| = 0$ . For notational convenience, we assume

<sup>3</sup>The condition  $f \leq m$  is required for solving the LS problem (30).

<sup>4</sup>Note that there are two errors in Proposition 5 of [18, p.5], where the isometry constant  $\delta_{(1+3\lceil 3/\rho^2 \rceil)K}$  and the iteration number  $\bar{n} = 2\lceil 3/\rho^2 \rceil |T \setminus T^0|$  should be replaced with  $\delta_{(1+7\lceil 3/\rho^2 \rceil)K}$  and  $\bar{n} = 4\lceil 3/\rho^2 \rceil |T \setminus T^0|$ , respectively. As a result, the condition for the  $\text{OMP}_{cK}$  (i.e., when  $\rho = 1$ ) become  $\delta_{22K} < 1/6$  with  $12K$  maximal iterations [24].

that  $\{x_i\}_{i=1,2,\dots,|\Gamma^k|}$  are arranged in descending order of their magnitudes, i.e.,

$$|x_1| \geq |x_2| \geq \dots \geq |x_{|\Gamma^k|}|. \quad (32)$$

Also, we define the subset  $\Gamma_\tau^k$  of  $\Gamma^k$  as

$$\Gamma_\tau^k = \begin{cases} \emptyset & \tau = 0, \\ \{1, 2, \dots, 2^{\tau-1}\} & \tau = 1, 2, \dots, \lceil \log_2 |\Gamma^k| \rceil, \\ \Gamma^k & \tau = \lceil \log_2 |\Gamma^k| \rceil + 1. \end{cases} \quad (33)$$

See Fig. 1 for the illustration of  $\Gamma_\tau^k$ . Notice that the last set  $\Gamma_{\lceil \log_2 |\Gamma^k| \rceil + 1}^k$  may have less than  $2^{\lceil \log_2 |\Gamma^k| \rceil}$  elements. For example, if  $\Gamma^k = \{1, 2, \dots, 7\}$ , then  $\Gamma_0^k = \emptyset$ ,  $\Gamma_1^k = \{1\}$ ,  $\Gamma_2^k = \{1, 2\}$ ,  $\Gamma_3^k = \{1, 2, 3, 4\}$ , and  $\Gamma_4^k = \{1, 2, \dots, 7\} = \Gamma^k$ .

**Lemma 3.3:** For a given set  $\Gamma^k$  and a constant  $\sigma \geq 2$ , let  $L_{k,\sigma} \in \{1, 2, \dots, \lceil \log_2 |\Gamma^k| \rceil + 1\}$  be the minimum positive integer satisfying<sup>5</sup>

$$\|\mathbf{x}_{\Gamma^k \setminus \Gamma_0^k}\|_2^2 < \sigma \|\mathbf{x}_{\Gamma^k \setminus \Gamma_1^k}\|_2^2, \quad (34)$$

$$\|\mathbf{x}_{\Gamma^k \setminus \Gamma_1^k}\|_2^2 < \sigma \|\mathbf{x}_{\Gamma^k \setminus \Gamma_2^k}\|_2^2, \quad (35)$$

⋮

$$\|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-2}^k}\|_2^2 < \sigma \|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2, \quad (36)$$

$$\|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2 \geq \sigma \|\mathbf{x}_{\Gamma^k \setminus \Gamma_L^k}\|_2^2. \quad (37)$$

Then,

- 1)  $|\Gamma_L^k| \leq |\Gamma^k|$ ,
- 2)  $|\Gamma_L^k| \leq 2^{L-1}$ ,
- 3)  $|\Gamma^k| \geq 2^{L-2}$ ,
- 4)  $|\Gamma^k| > \left(\frac{2\sigma-1}{2\sigma-2}\right) 2^{L-2}$ .

*Proof:* See Appendix B. ■

The following proposition provides a lower bound of the residual power difference ( $\|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2$ ) in the  $(l+1)$ -th ( $l \geq k$ ) iteration of  $\text{OMP}_{cK}$ .

**Proposition 3.4:** Let  $\mathbf{x} \in \mathcal{R}^n$  be any  $K$ -sparse vector supported on  $T$  and  $\Phi \in \mathcal{R}^{m \times n}$  be the sensing matrix with unit  $\ell_2$ -norm columns. Then, for a given  $\Gamma^k$  and an integer  $l \geq k$ , the residual of the  $\text{OMP}_{cK}$  satisfies

$$\|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2 \geq \frac{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}{|\Gamma_\tau^k|} (\|\mathbf{r}^l\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2) \quad (38)$$

where  $\tau = 0, 1, \dots, \lceil \log_2 |\Gamma^k| \rceil + 1$ .

*Proof:* See Appendix C. ■

From Proposition 3.4, we further obtain the following proposition, which plays an important role in the proof of Theorem 3.1.

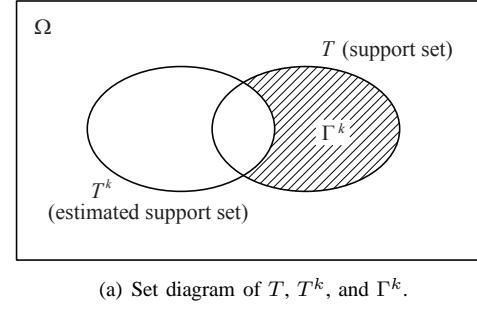
**Proposition 3.5:** For any integer  $l' > l$  and  $\tau \in \{0, 1, \dots, \lceil \log_2 |\Gamma^k| \rceil + 1\}$ , the residual of the  $\text{OMP}_{cK}$  satisfies

$$\|\mathbf{r}^{l'}\|_2^2 \leq C_{\tau, l, l'} \|\mathbf{r}^l\|_2^2 + (1 + \delta_{|\Gamma^k|}) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2, \quad (39)$$

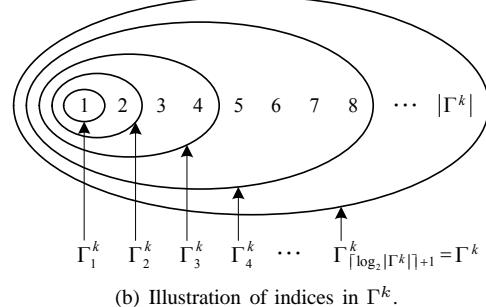
where

$$C_{\tau, l, l'} = \exp \left( -\frac{(l' - l)(1 - \delta_{|\Gamma_\tau^k \cup T^{l'-1}|})}{|\Gamma_\tau^k|} \right). \quad (40)$$

<sup>5</sup>In the sequel, we use  $L$  instead of  $L_{k,\sigma}$  for notation convenience.



(a) Set diagram of  $T$ ,  $T^k$ , and  $\Gamma^k$ .



(b) Illustration of indices in  $\Gamma^k$ .

Fig. 1. Illustration of sets  $T$ ,  $T^k$ , and  $\Gamma^k$ .

*Proof:* See Appendix D. ■

### B. Proof of Theorem 3.1

Our proof is based on the mathematical induction of the number of remaining indices  $|\Gamma^k|$  (after  $k$  iterations). We first check the case when  $|\Gamma^k| = 0$ . This case is trivial since it implies that all support indices have been selected (i.e.,  $T \subseteq T^k$ ) so that no more iteration is required. Next, we assume that the argument holds up to  $|\Gamma^k| = \gamma - 1$  ( $1 \leq \gamma < K$ ). In other words, we assume that if  $|\Gamma^k| = j$  ( $1 \leq j \leq \gamma - 1$ ), then maximally  $6j$  additional iterations are required to select all the support indices in  $\Gamma^k$ . Under this assumption, we show that if  $|\Gamma^k| = \gamma$ , it requires maximally  $6|\Gamma^k| = 6\gamma$  additional iterations to find out the rest of support indices (i.e.,  $f - k \leq 6\gamma$  where  $f$  is the final iteration number such that  $T \subseteq T^f$ ).

Although the details of this proof are somewhat cumbersome, the main idea is simple. When  $|\Gamma^k| = \gamma$ , we first show that the decent amount of indices in  $\Gamma^k$  can be selected within a specified number of additional iterations. To be specific, let  $k_i = k + \lceil 1.5(1 + \sum_{\tau=1}^i |\Gamma_\tau^k|) \rceil$  for  $i = 0, 1, \dots, L$  ( $\Gamma_\tau^k$  and  $L$  are defined in (33) and Lemma 3.3). Then, we show that under  $\delta_{|T \cup T^{k_L}|} < 0.03248$ , the  $\text{OMP}_{cK}$  algorithm selects more than  $2^{L-2}$  support indices in  $\Gamma^k$  within  $k_L - k$  additional iterations (i.e., from the  $(k+1)$ -th to the  $k_L$ -th iteration). In other word, the number of remaining support indices in  $\Gamma^{k_L}$  (after  $k_L$  iterations of the  $\text{OMP}_{cK}$ ) satisfies

$$|\Gamma^{k_L}| < \gamma - 2^{L-2}. \quad (41)$$

Second, we show that by the induction hypothesis, it requires no more than  $6|\Gamma^{k_L}|$  iterations to select the rest of support indices in  $\Gamma^{k_L}$  (i.e.,  $f - k_L \leq 6|\Gamma^{k_L}|$ ). In summary, the goal of the induction step is to show

$$k_L + 6(\gamma - 2^{L-2}) - k \leq 6\gamma. \quad (42)$$

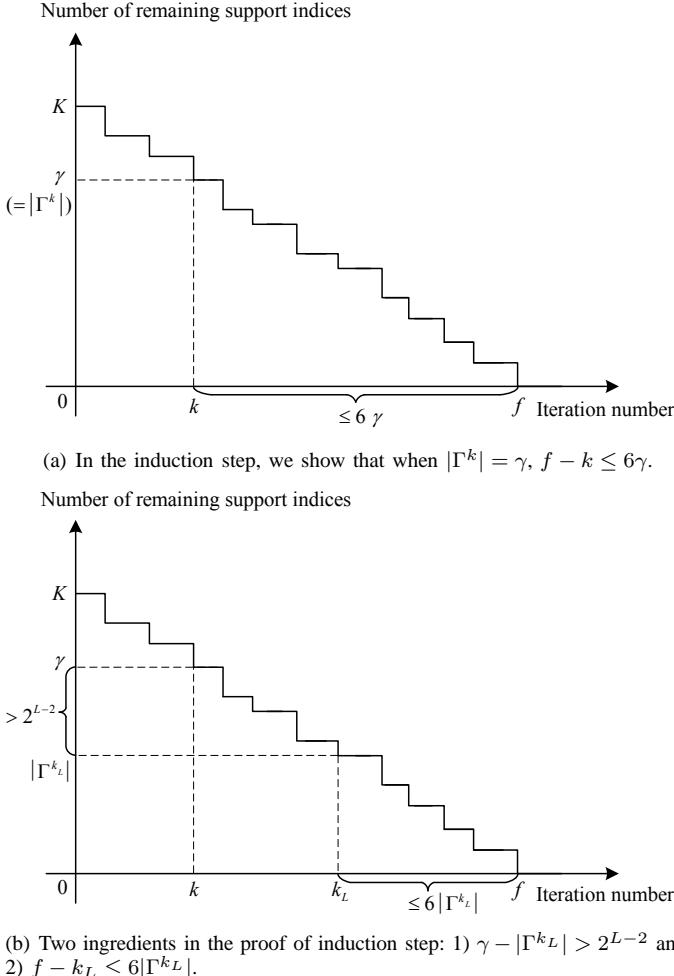


Fig. 2. Illustration of the induction step.

An illustration of the induction step is described in Fig. 2.

Before we proceed, we briefly explain the key steps to prove (41). Consider  $\mathbf{x}_{\Gamma^k_L}$  and  $\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}$ , which are two truncated vectors of  $\mathbf{x}_{\Gamma^k}$ . From the definition of  $\Gamma_\tau^k$  in (33), we have  $\Gamma_{L-1}^k = \{1, 2, \dots, 2^{L-2}\}$  and  $\Gamma^k \setminus \Gamma_{L-1}^k = \{2^{L-2}+1, \dots, \gamma\}$  so that an alternative form of (41) is

$$|\Gamma^k_L| < |\Gamma^k \setminus \Gamma_{L-1}^k|. \quad (43)$$

Further, recalling that  $\{x_i\}_{i=1,2,\dots,\gamma}$  are arranged in descending order of their magnitudes (i.e.,  $|x_1| \geq |x_2| \geq \dots \geq |x_\gamma|$ ), it is clear that  $\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}$  consists of  $\gamma - 2^{L-2}$  most non-significant elements in  $\mathbf{x}_{\Gamma^k}$ . Hence, one can easily check that the sufficient condition of (41) is

$$\|\mathbf{x}_{\Gamma^k_L}\|_2^2 < \|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2. \quad (44)$$

To show this, we construct upper and lower bounds of  $\|\mathbf{r}^{k_L}\|_2^2$  as

$$\|\mathbf{r}^{k_L}\|_2^2 \leq B_u \|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2, \quad (45)$$

$$\|\mathbf{r}^{k_L}\|_2^2 \geq B_\ell \|\mathbf{x}_{\Gamma^k_L}\|_2^2. \quad (46)$$

Since (44) holds true under  $B_u < B_\ell$  and also (44) is a sufficient condition of (41), it is clear that (41) can be guaranteed as long as  $B_u < B_\ell$ .

First, we find out an upper bound of  $\|\mathbf{r}^{k_L}\|_2^2$ . Since  $|\Gamma^k| = \gamma$ , by applying Proposition 3.5, we have

$$\|\mathbf{r}^{k_1}\|_2^2 \leq C_{1,k_0,k_1} \|\mathbf{r}^k\|_2^2 + (1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_1^k}\|_2^2, \quad (47)$$

$$\|\mathbf{r}^{k_2}\|_2^2 \leq C_{2,k_1,k_2} \|\mathbf{r}^{k_1}\|_2^2 + (1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_2^k}\|_2^2, \quad (48)$$

$$\vdots$$

$$\|\mathbf{r}^{k_L}\|_2^2 \leq C_{L,k_{L-1},k_L} \|\mathbf{r}^{k_{L-1}}\|_2^2 + (1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_L^k}\|_2^2. \quad (49)$$

Noting that  $k_i = k + \lceil 1.5(1 + \sum_{\tau=1}^i |\Gamma_\tau^k|) \rceil$  for  $i = 0, 1, \dots, L$ , it follows from (40) that<sup>6</sup>

$$C_{i,k_{i-1},k_i} = \begin{cases} \exp\left(-\frac{1.5(1 + |\Gamma_1^k|)(1 - \delta_{|\Gamma_1^k \cup T^{k_1-1}|})}{|\Gamma_1^k|}\right), & i = 1, \\ \exp\left(-\frac{1.5|\Gamma_i^k|(1 - \delta_{|\Gamma_i^k \cup T^{k_{i-1}-1}|})}{|\Gamma_i^k|}\right), & i = 2, \dots, L-1, \\ \exp\left(-\frac{1.5|\Gamma_L^k|(1 - \delta_{|\Gamma_L^k \cup T^{k_L-1}|})}{|\Gamma_L^k|}\right), & i = L. \end{cases} \quad (50)$$

One can easily show that for  $i = 1, \dots, L$ ,

$$C_{i,k_{i-1},k_i} \leq \exp\left(-\frac{1.5|\Gamma_i^k|(1 - \delta_{|\Gamma_i^k \cup T^{k_{i-1}-1}|})}{|\Gamma_i^k|}\right) \quad (51)$$

$$\leq \exp\left(-1.5(1 - \delta_{|\Gamma_i^k \cup T^{k_{i-1}-1}|})\right) \quad (52)$$

$$\leq \exp\left(-1.5(1 - \delta_{|\Gamma_L^k \cup T^{k_L-1}|})\right) \quad (53)$$

where (53) is from the monotonicity of isometry constant in Lemma 2.2.<sup>7</sup>

For notational simplicity, we let  $\eta = \exp(-1.5(1 - \delta_{|\Gamma_L^k \cup T^{k_L-1}|}))$ . Then (47)–(49) can be rewritten as

$$\|\mathbf{r}^{k_1}\|_2^2 \leq \eta \|\mathbf{r}^k\|_2^2 + (1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_1^k}\|_2^2, \quad (54)$$

$$\|\mathbf{r}^{k_2}\|_2^2 \leq \eta \|\mathbf{r}^{k_1}\|_2^2 + (1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_2^k}\|_2^2, \quad (55)$$

$$\vdots$$

$$\|\mathbf{r}^{k_L}\|_2^2 \leq \eta \|\mathbf{r}^{k_{L-1}}\|_2^2 + (1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_L^k}\|_2^2. \quad (56)$$

Multiplying (54) by  $\eta^{L-1}$  and (55) by  $\eta^{L-2}$ , and so on, and then summing these rows, we have

$$\|\mathbf{r}^{k_L}\|_2^2 \leq \eta^L \|\mathbf{r}^k\|_2^2 + (1 + \delta_\gamma) \sum_{\tau=1}^L \eta^{L-\tau} \|\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2. \quad (57)$$

In order to obtain an upper bound of  $\|\mathbf{r}^{k_L}\|_2^2$ , we need to find out an upper bound of  $\|\mathbf{r}^k\|_2^2$ .

<sup>6</sup>Since  $|\Gamma_i^k| = 2^{i-1}$  for  $i = 1, 2, \dots, L-1$ , we have  $k_1 - k_0 = \lceil 1.5(1 + |\Gamma_1^k|) \rceil = 1.5(1 + |\Gamma_1^k|)$  and  $k_i - k_{i-1} = k + \lceil 1.5(1 + \sum_{\tau=1}^i |\Gamma_\tau^k|) \rceil - (k + \lceil 1.5(1 + \sum_{\tau=1}^{i-1} |\Gamma_\tau^k|) \rceil) = \lceil 1.5(1 + \sum_{\tau=1}^i 2^{i-1}) \rceil - \lceil 1.5(1 + \sum_{\tau=1}^{i-1} 2^{i-1}) \rceil = \lceil 1.5 \cdot 2^i \rceil - \lceil 1.5 \cdot 2^{i-1} \rceil = 1.5 \cdot 2^i - 1.5 \cdot 2^{i-1} = 1.5 \cdot 2^{i-1} = 1.5 |\Gamma_i^k|$  for  $i = 2, \dots, L-1$ .

<sup>7</sup>Since  $\Gamma_i^k \subseteq \Gamma_L^k$  and  $T^{k_{i-1}-1} \subseteq T^{k_L-1}$ , we have  $\Gamma_i^k \cup T^{k_{i-1}-1} \subseteq \Gamma_L^k \cup T^{k_L-1}$  and  $|\Gamma_i^k \cup T^{k_{i-1}-1}| \leq |\Gamma_L^k \cup T^{k_L-1}|$ .

*Lemma 3.6:* The residual  $\mathbf{r}^k$  satisfies

$$\|\mathbf{r}^k\|_2^2 \leq (1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_0^k}\|_2^2. \quad (58)$$

*Proof:* See Appendix E.  $\blacksquare$

Plugging (58) into (57) yields

$$\|\mathbf{r}^{k_L}\|_2^2 \leq (1 + \delta_\gamma) \sum_{\tau=0}^L \eta^{L-\tau} \|\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2. \quad (59)$$

Also, from the definition of  $L$  in Lemma 3.3, we have

$$\|\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \leq \sigma^{L-1-\tau} \|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2 \quad (60)$$

for  $\tau = 1, 2, \dots, L$ . Applying (60) to (59) yields

$$\|\mathbf{r}^{k_L}\|_2^2 \leq \frac{(1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2}{\sigma} \sum_{\tau=0}^L (\sigma\eta)^{L-\tau} \quad (61)$$

$$= \frac{(1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2}{\sigma} \sum_{\tau=0}^L (\sigma\eta)^\tau. \quad (62)$$

When  $0 < \sigma\eta < 1$ , we further have

$$\|\mathbf{r}^{k_L}\|_2^2 < \frac{(1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2}{\sigma} \sum_{\tau=0}^{\infty} (\sigma\eta)^\tau \quad (63)$$

$$= \frac{(1 + \delta_\gamma)}{\sigma(1 - \sigma\eta)} \|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2, \quad (64)$$

which is the desired upper bound of  $\|\mathbf{r}^{k_L}\|_2^2$  (see (45)). That is,

$$B_u = \frac{1 + \delta_\gamma}{\sigma(1 - \sigma\eta)}. \quad (65)$$

Second, we build a lower bound of  $\|\mathbf{r}^{k_L}\|_2^2$ . Noting that  $\|\mathbf{x} - \mathbf{x}^{k_L}\|_0 \leq |T \cup T^{k_L}|$ , we have

$$\|\mathbf{r}^{k_L}\|_2^2 = \|\Phi(\mathbf{x} - \mathbf{x}^{k_L})\|_2^2 \quad (66)$$

$$\geq (1 - \delta_{|T \cup T^{k_L}|}) \|\mathbf{x} - \mathbf{x}^{k_L}\|_2^2 \quad (67)$$

$$\geq (1 - \delta_{|T \cup T^{k_L}|}) \|\mathbf{x}_{\Gamma^{k_L}}\|_2^2. \quad (68)$$

Thus,

$$B_\ell = 1 - \delta_{|T \cup T^{k_L}|}. \quad (69)$$

As mentioned, it suffices to show  $B_u < B_\ell$  to prove (41). Using (65) and (69), we have

$$B_u = \frac{1 + \delta_\gamma}{\sigma(1 - \sigma\eta)(1 - \delta_{|T \cup T^{k_L}|})} B_\ell. \quad (70)$$

Note that  $0 < \sigma\eta < 1$  and

$$\sigma(1 - \sigma\eta) = -\frac{1}{\eta} \left( \sigma\eta - \frac{1}{2} \right)^2 + \frac{1}{4\eta}. \quad (71)$$

Hence, by choosing  $\sigma\eta = \frac{1}{2}$ ,  $\sigma(1 - \sigma\eta)$  takes the maximum value  $\frac{1}{4\eta}$  and

$$B_u = \frac{4\eta(1 + \delta_\gamma)}{1 - \delta_{|T \cup T^{k_L}|}} B_\ell \quad (72)$$

$$= \frac{4(1 + \delta_\gamma) \exp(-1.5(1 - \delta_{|T \cup T^{k_L}|}))}{1 - \delta_{|T \cup T^{k_L}|}} B_\ell. \quad (73)$$

From the monotonicity of isometry constant, we have

$$|\Gamma^k| \leq |T \cup T^{k_L}| \Rightarrow \delta_\gamma \leq \delta_{|T \cup T^{k_L}|}, \quad (74)$$

$$|\Gamma_L^k \cup T^{k_L-1}| \leq |T \cup T^{k_L}| \Rightarrow \delta_{|\Gamma_L^k \cup T^{k_L-1}|} \leq \delta_{|T \cup T^{k_L}|}. \quad (75)$$

From (74), (75), and (73), we have

$$B_u \leq \frac{4(1 + \delta_{|T \cup T^{k_L}|}) \exp(-1.5(1 - \delta_{|T \cup T^{k_L}|}))}{1 - \delta_{|T \cup T^{k_L}|}} B_\ell. \quad (76)$$

When  $\delta_{|T \cup T^{k_L}|} < 0.03248$ , one can easily show that

$$\frac{4(1 + \delta_{|T \cup T^{k_L}|}) \exp(-1.5(1 - \delta_{|T \cup T^{k_L}|}))}{1 - \delta_{|T \cup T^{k_L}|}} < 1, \quad (77)$$

which implies  $B_u < B_\ell$  and hence we can conclude that (41) holds true under  $\delta_{|T \cup T^{k_L}|} < 0.03248$ .

We are now ready to prove the induction step. As shown in (41), the number of remaining support indices after  $k_L$  iterations satisfies  $|\Gamma^{k_L}| < \gamma - 2^{L-2}$ . Since  $\gamma - 2^{L-2} \leq \gamma - 1$ , we also have

$$|\Gamma^{k_L}| \leq \gamma - 1. \quad (78)$$

The induction hypothesis implies that all indices in  $\Gamma^{k_L}$  can be chosen within  $6|\Gamma^{k_L}|$  additional iterations of the OMP<sub>CK</sub>.<sup>8</sup> That is,

$$f - k_L \leq 6|\Gamma^{k_L}|. \quad (79)$$

Recall that our goal of the induction step is to show that it requires  $6\gamma$  maximal iterations to select all the support indices in  $\Gamma^k$ , i.e.,

$$f - k \leq 6\gamma. \quad (80)$$

From (79), it is clear that (80) holds true whenever

$$k_L + 6|\Gamma^{k_L}| - k \leq 6\gamma. \quad (81)$$

Using (41), we have

$$k_L + 6|\Gamma^{k_L}| - k < k_L + 6(\gamma - 2^{L-2}) - k. \quad (82)$$

Now, if we show that

$$k_L + 6(\gamma - 2^{L-2}) - k \leq 6\gamma, \quad (83)$$

then (80) holds true and the induction step is established.

From the definition of  $k_L$ ,

$$k_L = k + \lceil 1.5(1 + \sum_{\tau=1}^L |\Gamma_\tau^k|) \rceil \quad (84)$$

$$= k + \lceil 1.5(1 + \sum_{\tau=1}^{L-1} |\Gamma_\tau^k|) + 1.5|\Gamma_L^k| \rceil \quad (85)$$

$$= k + \lceil 1.5(1 + \sum_{\tau=1}^{L-1} 2^{\tau-1}) + 1.5|\Gamma_L^k| \rceil \quad (86)$$

$$= k + \lceil 1.5 \cdot 2^{L-1} + 1.5|\Gamma_L^k| \rceil, \quad (87)$$

<sup>8</sup>In the induction step, we assume that when  $|\Gamma^k| \leq \gamma - 1$ , it requires  $6|\Gamma^k|$  maximal iterations to find out all the indices in  $\Gamma^k$ .

where (86) uses the fact that  $|\Gamma_\tau^k| = 2^{\tau-1}$  for  $\tau = 1, 2, \dots, L-1$ . Then, it follows that

$$k_L + 6(\gamma - 2^{L-2}) - k = \lceil 1.5 \cdot 2^{L-1} + 1.5|\Gamma_L^k| \rceil + 6(\gamma - 2^{L-2}) \quad (88)$$

$$\leq \lceil 3 \cdot 2^{L-1} \rceil + 6(\gamma - 2^{L-2}) \quad (89)$$

$$= 3 \cdot 2^{L-1} + 6(\gamma - 2^{L-2}) \quad (90)$$

$$= 6\gamma, \quad (91)$$

where (89) is from  $|\Gamma_L^k| \leq 2^{L-1}$  in Lemma 3.3. This completes the induction.

In summary, we have

$$T^{k+6|\Gamma^k|} = T^{k+6\gamma} \quad (92)$$

$$\supseteq T^{k+(k_L+6(\gamma-2^{L-2})-k)} \quad (93)$$

$$\supseteq T^{k+(k_L+6|\Gamma^{k_L}|-k)} \quad (94)$$

$$= T^{k_L+6|\Gamma^{k_L}|} \quad (95)$$

$$\supseteq T^f \quad (96)$$

$$\supseteq T, \quad (97)$$

where (93) is from (91), (94) is due to (82), and (96) is from (79).

Thus far, we have shown that under  $\delta_{|T \cup T^{k_L}|} < 0.03248$ ,  $T \subseteq T^{k+6|\Gamma^k|}$ . We now need to find an order  $|T \cup T^{k_L}|$  of the isometry constant. In this step, we first show that  $|T \cup T^{k_L}| < k + \lceil 3.9\gamma \rceil$  and then show that  $k + \lceil 3.9\gamma \rceil \leq \lceil 8.93K \rceil$ . Combining these two inequalities, we obtain  $|T \cup T^{k_L}| \leq \lceil 8.93K \rceil$ .

First, observe that

$$|T \cup T^{k_L}| = |T \setminus T^{k_L}| + |T^{k_L}| \quad (98)$$

$$= |\Gamma^{k_L}| + |T^{k_L}| \quad (99)$$

$$< \gamma - 2^{L-2} + k_L, \quad (100)$$

where (100) follows from (41). Combining (87) and (100) yields

$$|T \cup T^{k_L}| < \gamma - 2^{L-2} + k + \lceil 1.5 \cdot 2^{L-1} + 1.5|\Gamma_L^k| \rceil \quad (101)$$

$$= \gamma - 2^{L-2} + k + \lceil 2^{L-1} + 2^{L-2} + 1.5|\Gamma_L^k| \rceil \quad (102)$$

$$= \gamma + 2^{L-1} + k + \lceil 1.5|\Gamma_L^k| \rceil \quad (103)$$

$$\leq k + 2^{L-1} + \lceil 2.5\gamma \rceil, \quad (104)$$

where (103) is because  $L \geq 1$  and hence  $1.5 \cdot 2^{L-1}$  is an integer and (104) is due to  $|\Gamma_L^k| \leq |\Gamma^k| = \gamma$  in Lemma 3.3.

Recalling that  $\sigma\eta = \frac{1}{2}$  and  $\eta = \exp(-1.5(1 - \delta_{|\Gamma_L^k \cup T^{k_L-1}|}))$ ,

$$\begin{aligned} \sigma &= 0.5 \exp\left(1.5(1 - \delta_{|\Gamma_L^k \cup T^{k_L-1}|})\right) \\ &< 0.5e^{1.5} \\ &= 2.2408. \end{aligned} \quad (105)$$

Using this together with property 4) in Lemma 3.3, we have

$$\begin{aligned} 2^{L-2} &< \left(\frac{2\sigma-2}{2\sigma-1}\right)|\Gamma^k| \\ &= \left(\frac{2\sigma-2}{2\sigma-1}\right)\gamma \\ &= \gamma - \frac{\gamma}{2\sigma-1} \\ &< 0.7128\gamma \end{aligned} \quad (106)$$

and

$$2^{L-1} \leq \lfloor 1.4256\gamma \rfloor. \quad (107)$$

Finally, using (104) and (107),

$$\begin{aligned} |T \cup T^{k_L}| &< k + 2^{L-1} + \lceil 2.5\gamma \rceil \\ &\leq k + \lfloor 1.4256\gamma \rfloor + \lceil 2.5\gamma \rceil \\ &\leq k + \lceil 3.9256\gamma \rceil \\ &< k + \lceil 3.93\gamma \rceil \end{aligned} \quad (108)$$

Next, we build an upper bound for  $k + \lceil 3.93\gamma \rceil$ . As mentioned,  $k$  is the number of iterations already performed and  $|\Gamma^k| = \gamma$  is the number of remaining correct indices. In order to find an upper bound for  $k + \lceil 3.93\gamma \rceil$ , the following three cases need to be considered.

- $k = 0$  (i.e., in the initial iteration): In this case, it is clear that  $\gamma = K$  and hence

$$k + \lceil 3.93\gamma \rceil = 3.93K. \quad (109)$$

- $0 < k \leq 5K$ : Since  $\gamma \leq K$ ,

$$k + \lceil 3.93\gamma \rceil \leq 5K + \lceil 3.93K \rceil = \lceil 8.93K \rceil. \quad (110)$$

It is worth mentioning that the case of “ $k = 5K$ ,  $\gamma = K$ ” corresponds to the worst case where the OMP<sub>cK</sub> does not select any support index for the first  $5K$  iterations. Since the required number of iterations cannot exceed  $6K$ , the remaining  $K$  support indices should be selected in the last  $K$  iterations.

- $5K < k \leq 6K$ : Again, since the required number of iterations for selecting  $K$  support indices cannot exceed  $6K$ , we must have  $\gamma \leq 6K - k$  and thus

$$k + \lceil 3.93\gamma \rceil \leq k + \lceil 3.93(6K - k) \rceil = \lceil 23.58K - 2.93k \rceil. \quad (111)$$

Since  $k > 5K$ , we further have

$$k + \lceil 3.93\gamma \rceil < \lceil 23.58K - 2.93 \cdot 5K \rceil = \lceil 8.93K \rceil. \quad (112)$$

In summary, for all cases

$$k + \lceil 3.93\gamma \rceil \leq \lceil 8.93K \rceil. \quad (113)$$

Using this together with (108), we have

$$|T \cup T^{k_L}| \leq \lceil 8.93K \rceil, \quad (114)$$

which completes the proof of Theorem 3.1.

#### IV. CONCLUSION

In this paper, we investigated the recovery condition of the OMP algorithm ensuring the perfect recovery of sparse signals for two distinct scenarios. In the first part of this paper, we have provided a simple bound of the OMP guaranteeing the exact recovery of  $K$ -sparse signals using  $K$ -iterations of the OMP algorithm. Our condition,  $\sqrt{K}\theta_{K,1} + \delta_K < 1$ , bridges the gap between MIP and RIP conditions and also embraces many of recently proposed optimal or near optimal bounds of the OMP algorithm [12], [14], [21].

In the second part of this paper, we have analyzed a recovery condition for the OMP algorithm when the iteration number is allowed to be more than  $K$ . We have shown that the OMP can perfectly recover any  $K$ -sparse signal within  $6K$  iterations under the RIP of order  $\lfloor 8.93K \rfloor$ , which improves the previous results ( $30K$  iterations in [17] and  $12K$  iterations in [18]). Considering that larger number of iterations results in higher computational complexity and also imposes on stricter limitation to the sparsity level  $K$ , the reduction on the iteration number offers practical benefits in complexity as well as the relaxation of the sparsity of underlying signals to be recovered.

#### APPENDIX A PROOF OF LEMMA 2.5

*Proof:* Let  $\mathbf{u} \in \mathcal{R}^{|I_1|}$  be a unit  $\ell_2$ -norm vector, then

$$\begin{aligned} & \max_{\mathbf{u}: \|\mathbf{u}\|_2=1} |\langle \Phi_{I_1} \mathbf{u}, \Phi_{I_2} \mathbf{x}_{I_2} \rangle| \\ &= \max_{\mathbf{u}: \|\mathbf{u}\|_2=1} \|\mathbf{u}'(\Phi'_{I_1} \Phi_{I_2} \mathbf{x}_{I_2})\|_2 \\ &= \|\Phi'_{I_1} \Phi_{I_2} \mathbf{x}_{I_2}\|_2, \end{aligned} \quad (115)$$

where the maximum of the inner product is achieved when  $\mathbf{u}$  is in the same direction of  $\Phi'_{I_1} \Phi_{I_2} \mathbf{x}_{I_2}$ , i.e.,

$$\mathbf{u} = \frac{\Phi'_{I_1} \Phi_{I_2} \mathbf{x}_{I_2}}{\|\Phi'_{I_1} \Phi_{I_2} \mathbf{x}_{I_2}\|_2}.$$

Moreover, by the definition of restricted orthogonality constant,

$$\begin{aligned} |\langle \Phi_{I_1} \mathbf{u}, \Phi_{I_2} \mathbf{x}_{I_2} \rangle| &\leq \theta_{|I_1|, |I_2|} \|\mathbf{u}\|_2 \|\mathbf{x}\|_2 \\ &= \theta_{|I_1|, |I_2|} \|\mathbf{x}\|_2. \end{aligned} \quad (116)$$

We obtain the desired result by combining (115) and (116).  $\blacksquare$

#### APPENDIX B PROOF OF LEMMA 3.3

*Proof:* The first three inequalities are trivial. The last inequality can be justified as follows. From (34), we have

$$\|\mathbf{x}_{\Gamma_{L-1}^k \setminus \Gamma_{L-2}^k}\|_2^2 < (\sigma - 1) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2. \quad (117)$$

Recalling from the definition of  $\Gamma_{\tau}^k$  and  $L$  (see (33) and Lemma 3.3),

$$\Gamma_{L-1}^k \setminus \Gamma_{L-2}^k = \{2^{L-3} + 1, \dots, 2^{L-2}\}, \quad (118)$$

$$\Gamma^k \setminus \Gamma_{L-1}^k = \{2^{L-2} + 1, \dots, |\Gamma^k|\}. \quad (119)$$

Noting that  $\{x_i\}_{i=1,2,\dots,|\Gamma^k|}$  are arranged in descending order of their magnitudes, the elements of  $\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}$  are  $|\Gamma^k| - 2^{L-2}$  most non-significant ones in  $\mathbf{x}_{\Gamma^k}$  (and hence smaller than the elements in  $\mathbf{x}_{\Gamma_{L-1}^k \setminus \Gamma_{L-2}^k}$ ). Since  $\sigma - 1 \geq 1$ , we obtain from (117) that

$$|\Gamma_{L-1}^k \setminus \Gamma_{L-2}^k| < (\sigma - 1) |\Gamma^k \setminus \Gamma_{L-1}^k|. \quad (120)$$

That is,

$$2^{L-3} < (\sigma - 1) (|\Gamma^k| - 2^{L-2}), \quad (121)$$

which completes the proof.  $\blacksquare$

#### APPENDIX C PROOF OF PROPOSITION 3.4

The following lemma is useful in proving Proposition 3.4.

*Lemma C.1:* Let  $\Delta \mathbf{r} = \mathbf{r}^l - \mathbf{r}^{l+1}$ . Then the OMP<sub>cK</sub> satisfies the following equalities.

- 1)  $\Delta \mathbf{r} = \mathbf{P}_{T^{l+1}} \mathbf{r}^l$ ,
- 2)  $\Delta \mathbf{r} = \mathbf{P}_{T^l}^{\perp} \Phi \mathbf{x}^{l+1}$ ,
- 3)  $\Delta \mathbf{r} = \mathbf{P}_{T^l}^{\perp} \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1}$ ,
- 4)  $\|\Delta \mathbf{r}\|_2 = \sqrt{\langle \mathbf{r}^l, \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \rangle}$ .

*Proof:*

- 1) Since  $\mathbf{r}^{l+1} = \mathbf{y} - \mathbf{P}_{T^{l+1}} \mathbf{y} = \mathbf{P}_{T^{l+1}}^{\perp} \mathbf{y}$ , we have

$$\Delta \mathbf{r} = \mathbf{r}^l - \mathbf{P}_{T^{l+1}}^{\perp} \mathbf{y} \quad (122)$$

$$= \mathbf{r}^l - \mathbf{P}_{T^{l+1}}^{\perp} (\mathbf{r}^l + \Phi \mathbf{x}^l) \quad (123)$$

$$= \mathbf{r}^l - \mathbf{P}_{T^{l+1}}^{\perp} \mathbf{r}^l \quad (124)$$

$$= \mathbf{P}_{T^{l+1}} \mathbf{r}^l \quad (125)$$

where (123) follows from  $\mathbf{r}^l = \mathbf{y} - \Phi \mathbf{x}^l$ , (124) uses the facts that  $\Phi \mathbf{x}^l = \mathbf{P}_{T^l} \mathbf{y}$  lies in  $\text{span}(\Phi_{T^{l+1}})$  and that  $\mathbf{P}_{T^{l+1}}^{\perp}$  cancels all components in  $\text{span}(\Phi_{T^{l+1}})$ .

- 2) From  $\mathbf{r}^{l+1} = \mathbf{y} - \Phi \mathbf{x}^{l+1}$ , we have

$$\mathbf{P}_{T^l}^{\perp} \Phi \mathbf{x}^{l+1} = \mathbf{P}_{T^l}^{\perp} \mathbf{y} - \mathbf{P}_{T^l}^{\perp} \mathbf{r}^{l+1} \quad (126)$$

$$= \mathbf{r}^l - \mathbf{r}^{l+1} + \mathbf{P}_{T^l} \mathbf{r}^{l+1} \quad (127)$$

$$= \Delta \mathbf{r}, \quad (128)$$

where (128) is because  $\mathbf{r}^{l+1}$  is orthogonal to  $\text{span}(\Phi_{T^l})$  (i.e.,  $\mathbf{P}_{T^l} \mathbf{r}^{l+1} = \mathbf{0}$ ).

- 3) Noting that  $\Phi \mathbf{x}^{l+1} = \mathbf{P}_{T^l} \mathbf{x}_{T^l}^{l+1} + \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1}$ , we have

$$\Delta \mathbf{r} = \mathbf{P}_{T^l}^{\perp} \Phi \mathbf{x}^{l+1} \quad (129)$$

$$= \mathbf{P}_{T^l}^{\perp} (\Phi_{T^l} \mathbf{x}_{T^l}^{l+1} + \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1}) \quad (130)$$

$$= \mathbf{P}_{T^l}^{\perp} \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1}, \quad (131)$$

where (131) is from the fact that  $\mathbf{P}_{T^l}^{\perp}$  cancels all components in  $\text{span}(\Phi_{T^l})$  (i.e.,  $\mathbf{P}_{T^l}^{\perp} \Phi_{T^l} = \mathbf{0}$ ).

- 4) Since  $\phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1}$  lies in  $\text{span}(\phi_{t^{l+1}})$ , we have

$$\mathbf{P}_{t^{l+1}} \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} = \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \quad (132)$$

and

$$\langle \mathbf{r}^l, \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \rangle = \langle \mathbf{r}^l, \mathbf{P}_{t^{l+1}} \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \rangle \quad (133)$$

$$= \langle \mathbf{P}_{t^{l+1}} \mathbf{r}^l, \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \rangle \quad (134)$$

$$= \langle \Delta \mathbf{r}, \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \rangle, \quad (135)$$

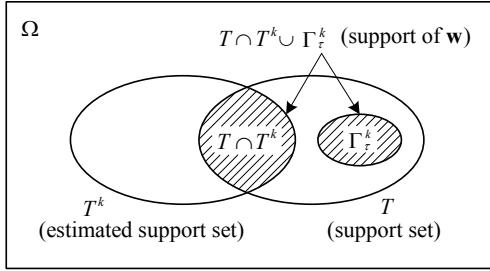


Fig. 3. Illustration of sets  $T$ ,  $T^k$ , and  $supp(\mathbf{w})$ .

where (134) is because  $\mathbf{P}_{t^{k+1}} = (\mathbf{P}_{t^{l+1}})'$  and (135) is due to (125). Using  $\mathbf{I} = \mathbf{P}_{T^l}^\perp + \mathbf{P}_{T^l}$ , we further have

$$\begin{aligned} & \langle \mathbf{r}^l, \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \rangle \\ &= \langle \Delta \mathbf{r}, \mathbf{P}_{T^l}^\perp \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \rangle + \langle \Delta \mathbf{r}, \mathbf{P}_{T^l} \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \rangle \end{aligned} \quad (136)$$

$$\begin{aligned} &= \langle \Delta \mathbf{r}, \Delta \mathbf{r} \rangle \\ &= \|\Delta \mathbf{r}\|_2^2, \end{aligned} \quad (137)$$

where (136) uses the fact that the vector  $\Delta \mathbf{r}$  is orthogonal to  $span(\mathbf{P}_{T^l})$  and hence  $\langle \Delta \mathbf{r}, \mathbf{P}_{T^l} \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \rangle = 0$ . This completes the proof.  $\blacksquare$

We are now ready to prove Proposition 3.4.

*Proof:* From Lemma C.1, we have

$$\|\Delta \mathbf{r}\|_2 = \frac{\langle \mathbf{r}^l, \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \rangle}{\|\Delta \mathbf{r}\|_2} \quad (138)$$

$$= \frac{\langle \mathbf{r}^l, \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1} \rangle}{\|\mathbf{P}_{T^l}^\perp \phi_{t^{l+1}} \mathbf{x}_{t^{l+1}}^{l+1}\|_2} \quad (139)$$

$$= \frac{\langle \phi_{t^{l+1}}, \mathbf{r}^l \rangle}{\|\mathbf{P}_{T^l}^\perp \phi_{t^{l+1}}\|_2}. \quad (140)$$

Noting that  $\|\mathbf{P}_{T^l}^\perp \phi_{t^{l+1}}\|_2 \leq \|\phi_{t^{l+1}}\|_2 = 1$  ( $\Phi$  has unit  $\ell_2$ -norm columns) and  $t^{l+1}$  corresponds to the maximum correlation between  $\mathbf{r}^l$  and the columns in  $\Phi$ , we have

$$\|\Delta \mathbf{r}\|_2 \geq \langle \phi_{t^{l+1}}, \mathbf{r}^l \rangle = \|\Phi' \mathbf{r}^l\|_\infty. \quad (141)$$

Since  $\mathbf{r}^l$  is orthogonal to  $span(\mathbf{P}_{T^l})$ ,  $\Phi'_{T^l} \mathbf{r}^l = \mathbf{0}$  and hence

$$\|\Phi' \mathbf{r}^l\|_\infty = \|(\Phi' \mathbf{r}^l)_{\Omega \setminus T^l}\|_\infty. \quad (142)$$

For a given  $\Gamma^k$ , let  $\Gamma_\tau^k \subseteq \Gamma^k$  where  $\tau \in \{1, 2, \dots, \lceil \log_2 |\Gamma^k| \rceil + 1\}$  be the set defined in (33). We define a sparse vector  $\mathbf{w} \in \mathcal{R}^n$  which equals  $\mathbf{x}$  for elements indexed by  $T \cap T^k \cup \Gamma_\tau^k$  (i.e.,  $\mathbf{w}_{T \cap T^k \cup \Gamma_\tau^k} = \mathbf{x}_{T \cap T^k \cup \Gamma_\tau^k}$ ) and  $\mathbf{0}$  otherwise. We consider the correlation  $\langle (\Phi' \mathbf{r}^l)_{\Omega \setminus T^l}, (\mathbf{w} - \mathbf{x}^l)_{\Omega \setminus T^l} \rangle$ . From Hölder's inequality,

$$\begin{aligned} & \langle (\Phi' \mathbf{r}^l)_{\Omega \setminus T^l}, (\mathbf{w} - \mathbf{x}^l)_{\Omega \setminus T^l} \rangle \\ & \leq \|(\Phi' \mathbf{r}^l)_{\Omega \setminus T^l}\|_\infty \|(\mathbf{w} - \mathbf{x}^l)_{\Omega \setminus T^l}\|_1. \end{aligned} \quad (143)$$

Since  $\mathbf{x}^l$  is supported on  $T^l$ , we have  $(\mathbf{x}^l)_{\Omega \setminus T^l} = \mathbf{0}$  and hence

$$\begin{aligned} & \|(\Phi' \mathbf{r}^l)_{\Omega \setminus T^l}\|_\infty \\ & \geq \frac{\langle (\Phi' \mathbf{r}^l)_{\Omega \setminus T^l}, (\mathbf{w} - \mathbf{x}^l)_{\Omega \setminus T^l} \rangle}{\|(\mathbf{w} - \mathbf{x}^l)_{\Omega \setminus T^l}\|_1} \end{aligned} \quad (144)$$

$$= \frac{\langle (\Phi' \mathbf{r}^l)_{\Omega \setminus T^l}, (\mathbf{w} - \mathbf{x}^l)_{\Omega \setminus T^l} \rangle}{\|\mathbf{w}_{\Omega \setminus T^l}\|_1} \quad (145)$$

$$= \frac{\langle \Phi' \mathbf{r}^l, \mathbf{w} - \mathbf{x}^l \rangle}{\|\mathbf{w}_{\Omega \setminus T^l}\|_1} \quad (146)$$

$$\geq \frac{\langle \Phi' \mathbf{r}^l, \mathbf{w} - \mathbf{x}^l \rangle}{\sqrt{\|\mathbf{w}_{\Omega \setminus T^l}\|_0} \|\mathbf{w}_{\Omega \setminus T^l}\|_2}, \quad (147)$$

where (146) is due to  $(\Phi' \mathbf{r}^l)_{T^l} = \mathbf{0}$  and (147) follows from the norm inequality  $\|\mathbf{v}\|_1 \leq \sqrt{\|\mathbf{v}\|_0} \|\mathbf{v}\|_2$ . Noting that  $supp(\mathbf{w}) = T \cap T^k \cup \Gamma_\tau^k$  and  $T^k \subseteq T^l$ , we have  $supp(\mathbf{w}_{\Omega \setminus T^l}) = (T \cap T^k \cup \Gamma_\tau^k) \setminus T^l = \Gamma_\tau^k \setminus T^l$  and hence

$$\|\mathbf{w}_{\Omega \setminus T^l}\|_0 = |\Gamma_\tau^k \setminus T^l| \leq |\Gamma_\tau^k|. \quad (148)$$

Applying (148) to (147),

$$\|(\Phi' \mathbf{r}^l)_{\Omega \setminus T^l}\|_\infty \geq \frac{\langle \Phi' \mathbf{r}^l, \mathbf{w} - \mathbf{x}^l \rangle}{\sqrt{|\Gamma_\tau^k|} \|\mathbf{w}_{\Omega \setminus T^l}\|_2}. \quad (149)$$

Combining (141), (142), and (149) yields

$$\|\Delta \mathbf{r}\|_2 \geq \frac{\langle \Phi' \mathbf{r}^l, \mathbf{w} - \mathbf{x}^l \rangle}{\sqrt{|\Gamma_\tau^k|} \|\mathbf{w}_{\Omega \setminus T^l}\|_2}. \quad (150)$$

Furthermore, we observe that

$$\begin{aligned} & \langle \Phi' \mathbf{r}^l, \mathbf{w} - \mathbf{x}^l \rangle \\ &= \langle \Phi(\mathbf{w} - \mathbf{x}^l), \mathbf{r}^l \rangle \\ &= \frac{1}{2} (\|\Phi(\mathbf{w} - \mathbf{x}^l)\|_2^2 + \|\mathbf{r}^l\|_2^2 - \|\Phi(\mathbf{w} - \mathbf{x}^l) - \mathbf{r}^l\|_2^2) \end{aligned} \quad (151)$$

$$\begin{aligned} &= \frac{1}{2} (\|\Phi(\mathbf{w} - \mathbf{x}^l)\|_2^2 + \|\mathbf{r}^l\|_2^2 - \|\Phi(\mathbf{w} - \mathbf{x})\|_2^2) \\ &= \frac{1}{2} (\|\Phi(\mathbf{w} - \mathbf{x}^l)\|_2^2 + \|\mathbf{r}^l\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2), \end{aligned} \quad (152)$$

where (151) uses the fact that  $\Phi \mathbf{x}^l + \mathbf{r}^l = \mathbf{y} = \Phi \mathbf{x}$ . When  $\|\mathbf{r}^l\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \geq 0$ , we apply the inequality  $\frac{1}{2}(a+b) \geq \sqrt{ab}$  to (152) to get

$$\begin{aligned} & \langle \Phi' \mathbf{r}^l, \mathbf{w} - \mathbf{x}^l \rangle \\ & \geq \|\Phi(\mathbf{w} - \mathbf{x}^l)\|_2 \sqrt{\|\mathbf{r}^l\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2}. \end{aligned} \quad (153)$$

Noting that  $\mathbf{w} - \mathbf{x}^l$  is supported on  $\Gamma_\tau^k \cup T^l$ , we have

$$\begin{aligned} & \|\Phi(\mathbf{w} - \mathbf{x}^l)\|_2 \\ & \geq \sqrt{1 - \delta_{|\Gamma_\tau^k \cup T^l|}} \|\mathbf{w} - \mathbf{x}^l\|_2 \end{aligned} \quad (154)$$

$$\geq \sqrt{1 - \delta_{|\Gamma_\tau^k \cup T^l|}} \|(\mathbf{w} - \mathbf{x}^l)_{\Omega \setminus T^l}\|_2 \quad (155)$$

$$= \sqrt{1 - \delta_{|\Gamma_\tau^k \cup T^l|}} \|\mathbf{w}_{\Omega \setminus T^l}\|_2, \quad (156)$$

where (156) is due to  $(\mathbf{x}^l)_{\Omega \setminus T^l} = \mathbf{0}$ . Plugging (153) and (156) into (150), we have

$$\|\Delta \mathbf{r}\|_2 \geq \sqrt{\frac{(1 - \delta_{|\Gamma_\tau^k \cup T^l|}) (\|\mathbf{r}^l\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2)}{|\Gamma_\tau^k|}}. \quad (157)$$

Further, since  $\|\Delta \mathbf{r}\|_2^2 = \|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2$ , we have

$$\begin{aligned} & \|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2 \\ & \geq \frac{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}{|\Gamma_\tau^k|} (\|\mathbf{r}^l\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2). \end{aligned} \quad (158)$$

Next, when  $\|\mathbf{r}^l\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 < 0$ , (158) holds naturally since  $\|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2 \geq 0$  (the residual is always non-increasing due to the orthogonal projection), which establishes the proof.  $\blacksquare$

## APPENDIX D PROOF OF PROPOSITION 3.5

*Proof:* Subtracting both sides of (38) by  $\|\mathbf{r}^l\|_2^2 - \|\Phi_{T \setminus \Gamma_\tau^k} \mathbf{x}_{T \setminus \Gamma_\tau^k}\|_2^2$ , we have

$$\begin{aligned} & \|\mathbf{r}^{l+1}\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \\ & \leq \left(1 - \frac{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}{|\Gamma_\tau^k|}\right) (\|\mathbf{r}^l\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2). \end{aligned} \quad (159)$$

Noting that

$$1 - \frac{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}{|\Gamma_\tau^k|} \leq \exp\left(-\frac{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}{|\Gamma_\tau^k|}\right),$$

we have

$$\begin{aligned} & \|\mathbf{r}^{l+1}\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \\ & \leq \exp\left(-\frac{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}{|\Gamma_\tau^k|}\right) (\|\mathbf{r}^l\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2) \end{aligned} \quad (160)$$

and also

$$\begin{aligned} & \|\mathbf{r}^{l+2}\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \\ & \leq \exp\left(-\frac{1 - \delta_{|\Gamma_\tau^k \cup T^{l+1}|}}{|\Gamma_\tau^k|}\right) \\ & \quad \times (\|\mathbf{r}^{l+1}\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2), \end{aligned} \quad (161)$$

⋮

$$\begin{aligned} & \|\mathbf{r}^{l'}\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \\ & \leq \exp\left(-\frac{1 - \delta_{|\Gamma_\tau^k \cup T^{l'-1}|}}{|\Gamma_\tau^k|}\right) \\ & \quad \times (\|\mathbf{r}^{l'-1}\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2). \end{aligned} \quad (162)$$

By multiplexing each row, we obtain

$$\begin{aligned} & \|\mathbf{r}^{l'}\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \\ & \leq \prod_{i=l}^{l'-1} \exp\left(-\frac{1 - \delta_{|\Gamma_\tau^k \cup T^i|}}{|\Gamma_\tau^k|}\right) \\ & \quad \times (\|\mathbf{r}^l\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2). \end{aligned} \quad (163)$$

Since  $\delta_{|\Gamma_\tau^k \cup T^l|} \leq \delta_{|\Gamma_\tau^k \cup T^{l+1}|} \leq \dots \leq \delta_{|\Gamma_\tau^k \cup T^{l'-1}|}$ , we further have

$$\begin{aligned} & \|\mathbf{r}^{l'}\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \\ & \leq C_{\tau, l, l'} (\|\mathbf{r}^l\|_2^2 - \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2), \end{aligned} \quad (164)$$

where

$$C_{\tau, l, l'} = \exp\left(-\frac{(l' - l)(\delta_{|\Gamma_\tau^k \cup T^{l'-1}|})}{|\Gamma_\tau^k|}\right). \quad (165)$$

Noting that  $C_{\tau, l, l'} > 0$ , we derive from (164) that

$$\|\mathbf{r}^{l'}\|_2^2 \leq C_{\tau, l, l'} \|\mathbf{r}^l\|_2^2 + \|\Phi_{\Gamma^k \setminus \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2. \quad (166)$$

Furthermore, since  $|\Gamma^k \setminus \Gamma_\tau^k| \leq |\Gamma^k|$ , the RIP implies that

$$\|\mathbf{r}^{l'}\|_2^2 \leq C_{\tau, l, l'} \|\mathbf{r}^l\|_2^2 + (1 + \delta_{|\Gamma^k|}) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2, \quad (167)$$

which completes the proof.  $\blacksquare$

## APPENDIX E PROOF OF LEMMA 3.6

*Proof:*  $T^k \cap T \subseteq T^k$  implies that

$$\|\mathbf{r}^k\|_2^2 = \|\mathbf{P}_{T^k}^\perp \mathbf{y}\|_2^2 \leq \|\mathbf{P}_{T^k \cap T}^\perp \mathbf{y}\|_2^2. \quad (168)$$

Also, noting that  $\mathbf{P}_{T^k \cap T}^\perp \mathbf{y}$  is the projection of  $\mathbf{y}$  onto the orthogonal complement of  $\text{span}(\Phi_{T^k \cap T})$ ,

$$\|\mathbf{P}_{T^k \cap T}^\perp \mathbf{y}\|_2^2 = \min_{\text{supp}(\mathbf{v})=T^k \cap T} \|\mathbf{y} - \Phi \mathbf{v}\|_2^2. \quad (169)$$

Thus,

$$\|\mathbf{r}^k\|_2^2 \leq \|\mathbf{y} - \Phi_{T^k \cap T} \mathbf{x}_{T^k \cap T}\|_2^2 \quad (170)$$

$$= \|\Phi_{T^k} \mathbf{x}_T - \Phi_{T^k \cap T} \mathbf{x}_{T^k \cap T}\|_2^2 \quad (171)$$

$$= \|\Phi_{\Gamma^k} \mathbf{x}_{\Gamma^k}\|_2^2 \quad (172)$$

$$\leq (1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k}\|_2^2 \quad (173)$$

$$= (1 + \delta_\gamma) \|\mathbf{x}_{\Gamma^k \setminus \Gamma_0^k}\|_2^2, \quad (174)$$

where (172) is from  $T \setminus (T^k \cap T) = T \setminus T^k = \Gamma^k$ , (173) is from the definition of RIP where  $|\Gamma^k| = \gamma$ , and (174) is due to  $\Gamma_0^k = \emptyset$ .  $\blacksquare$

## REFERENCES

- [1] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decompositions," *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.
- [2] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [3] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [4] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [5] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [6] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 2210–2219, May 2008.
- [7] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inform. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

- [8] T. Zhang, "Adaptive forward-backward greedy algorithm for learning sparse representations," *IEEE Trans. Inform. Theory*, vol. 57, no. 7, pp. 4689–4708, Jul. 2011.
- [9] T. T. Cai, L. Wang, and G. Xu, "Shifting inequality and recovery of sparse signals," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1300–1308, Mar. 2010.
- [10] M. A. Davenport and M. B. Wakin, "Analysis of Orthogonal Matching Pursuit using the restricted isometry property," *IEEE Trans. Inform. Theory*, vol. 56, no. 9, pp. 4395–4401, Sep. 2010.
- [11] S. Foucart, "Sparse recovery algorithms: sufficient conditions in terms of restricted isometry constants," *Approximation Theory XIII: San Antonio 2010*, pp. 65–77, 2012.
- [12] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
- [13] E. Liu and V. N. Temlyakov, "The orthogonal super greedy algorithm and applications in compressed sensing," *IEEE Trans. Inform. Theory*, vol. 58, no. 4, pp. 2040–2047, Apr. 2012.
- [14] J. Wang and B. Shim, "On the recovery limit of orthogonal matching pursuit using restricted isometric property," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4973–4976, Sep. 2012.
- [15] Q. Mo and Y. Shen, "A remark on the restricted isometry property in orthogonal matching pursuit algorithm," *IEEE Trans. Inform. Theory*, vol. 58, no. 6, pp. 3654–3656, Jun. 2012.
- [16] E. Livshitz, "On efficiency of orthogonal matching pursuit," 2010 [Online]. Available: [http://arxiv.org/PS\\_cache/arxiv/pdf/1004/1004.3946v1.pdf](http://arxiv.org/PS_cache/arxiv/pdf/1004/1004.3946v1.pdf).
- [17] T. Zhang, "Sparse recovery with orthogonal matching pursuit under rip," *IEEE Trans. Inform. Theory*, vol. 57, no. 9, pp. 6215–6221, Sep. 2011.
- [18] S. Foucart, "Stability and robustness of weak orthogonal matching pursuits," 2011 [Online]. Available: <http://www.math.drexel.edu/~foucart/WOMMP.pdf>.
- [19] J. Wang, S. Kwon, and B. Shim, "Generalized orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 60, no. 12, Dec. 2012.
- [20] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via  $\ell_1$ -minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197, 2003.
- [21] J. Wang, S. Kwon, and B. Shim, "Near optimal bound of orthogonal matching pursuit using restricted isometric constant," *EURASIP J Adv. Signal Process.*, vol. 2012, no. 1, pp. 8, 2012.
- [22] D. Needell and J. A. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, Mar. 2009.
- [23] T. T. Cai, L. Wang, and G. Xu, "Stable recovery of sparse signals and an oracle inequality," *IEEE Trans. Inform. Theory*, vol. 56, no. 7, pp. 3516–3522, Jul. 2010.
- [24] J. Wang, "Private communication with S. Foucart," 2012.